

## THE 1991 NATIONAL CENSUS TEST

Michael J. Sheridan and Scott Buchanan, Statistics Canada  
Michael J. Sheridan, 6-A2 Jean Talon Building, Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0T6

**KEY WORDS:** questionnaire design, qualitative research, user friendly, content development, respondent error.

### 1. Introduction

In November, 1988, Statistics Canada conducted the first of two National Census Tests. The National Census Test (N.C.T.) represented for Statistics Canada, the culmination of three major research endeavors in the development of the 1991 Census of Canada. These components were a national consultation program, two modular tests and a series of qualitative research tasks.

The principal objectives established for the National Census Test were to detect and evaluate response error, establish on a question by question basis the field edit failure rates, and assess, to the degree possible, the Canadian public's reaction to the content of the questionnaire (i.e. time to complete, sensitivity of questions, etc.). It should be noted that the N.C.T. was not intended as a dress rehearsal for the 1991 Census, but rather as a content test.

This paper will discuss and assess the questionnaire testing program for the 1991 Census of Canada in terms of a review of the model that was used in the testing program, as well as a description of the component parts of the testing model. These major components are presented as a review of how the consultation program, the modular tests and the qualitative research conducted as inputs to the National Census Test, impacted on its development. The paper presents an overview of the methodology used for the N.C.T. as well as focusing on how the National Census Test contributed to the development of a more "user friendly" census questionnaire. The paper also presents a review of several analytical techniques on key variables as an indication of the analysis components that are used to evaluate and assess the test results.

### 2. Important Inputs to the National Census Test

Critical to the development process of the content of the N.C.T. were a series of consultations which took place during the 12 to 16 months prior to the tests. Throughout 1987, extensive consultations were conducted on the content of the 1991 Census. Consultations for past Censuses were generally limited to major users of Census data. However, in the consultation process for the 1991 Census of Canada the scope of the consultations was significantly broadened and extended to include virtually every sector of Canadian Society including the general public. In total, 88 meetings were held in almost every major city in Canada. Over 1,200 people participated in the process with more than 150 written briefs being submitted. The end result of the consultation was the development of 46 key recommendations for the 1991 Census content. These recommendations were one of the significant factors in the targeting of testing some of the new content items developed for the National Census Test, as well as sources for suggesting changes to other "standing" content items.

The second component in the content development model were the results of two modular tests. The modular tests were conducted in late 1987 and early 1988 using non-probability samples of approximately 3,000 households.

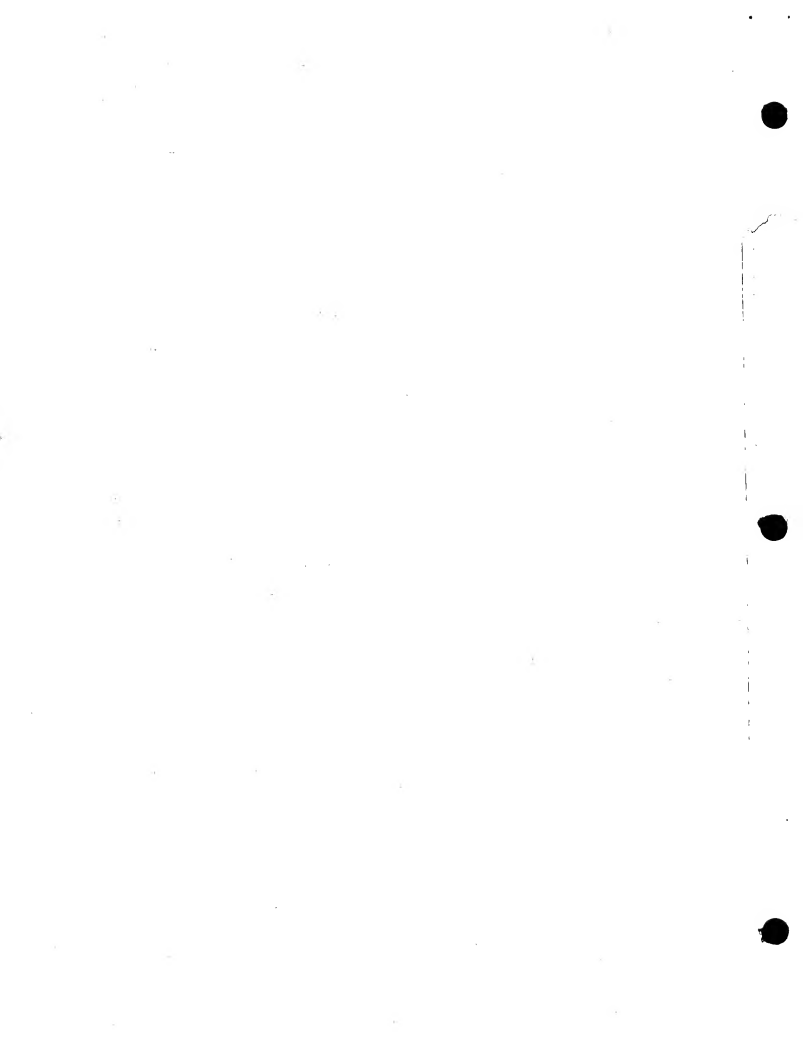
These tests are referred to as MT1 and MT2. MT1 had, as its focus, the revisions to previously used nuptiality and fertility questions as well as some new labour market and income questions. Modular Test 2 focussed on language, ethno-cultural and mobility status questions. The modular tests were of significant importance to the N.C.T. project because they served as a "pre-testing" ground for a number of proposed question reformulations and designs. These modular tests really served as a filtering mechanism, in that the analysis of the results defined, in a number of instances, questions which did not work. Consequently, the project charged with the development of the N.C.T. used the results to establish those viable question options which could, or should, be included in the major National test.

One of the key components in the model used for the development of the content for the National Census Test was the use of qualitative research. The qualitative research endeavours were divided into two parts and involved the use of a number of focus groups and other qualitative research methods including "think aloud interviews" and interview observation. In the first stage of this qualitative research, Statistics Canada contracted the private sector research firm of Price Waterhouse to conduct an assessment of the questionnaire used in the 1986 Census. In terms of meeting the objective of developing a more respondent friendly questionnaire, the results of this qualitative research on the 1986 Census form was invaluable. The first phase of the research addressed the following questions with respect to the 1986 Census form:

1. What problems do respondents have in completing the form?
2. Are the instructions read?
3. Is the questionnaire accurately completed?
4. Is the format easy to follow?

The second stage of the qualitative research was conducted after the design and layout of the N.C.T. questionnaire but prior to its field testing. The reason for this was to ensure that the design, wording and layout of the questionnaire to be used in the test conformed, to the degree possible, to the recommendations emanating from the first evaluation.

The results of the Price Waterhouse research pointed out the fact that the 1986 Census questionnaire had major formatting problems that caused respondents to either become frustrated with the form or to complete it incorrectly. Some of these problems included lack of clear instructions as to where to begin the form and how to proceed with the completion of the questionnaire. The 1986 short form contained a flap page which caused many respondents to overlook the first two questions on the form. Many respondents encountered difficulties with what were thought to be logical skip patterns. In addition, it was found that the special instructions or notes on the questionnaire were either never read, or when they were read they tended to confuse rather than clarify. In general, it was felt that there was too



much verbiage on the form. Additional analysis also indicated that the level of the language utilized in the questionnaire was too complex for respondents. The results of the reports from this qualitative research provided a good starting point for attempted design improvements to the N.C.T. questionnaire. A description of the attempted improvements in the N.C.T. questionnaire are presented in section 8 of this paper.

### 3. Methodology of the National Census Test

The National Census Test sample was based on the design of the Monthly Canadian Labour Force Survey which provides for Canada ostensibly the equivalent measures from the U.S. Current Population Survey (C.P.S.). The Labour Force Survey is a monthly survey whose sample of individuals is representative of the civilian non-institutionalized population, 15 years of age or older, who live in Canada's ten provinces. The Labour Force Survey (L.F.S.) is based on a frame where the dwelling is the final unit of selection. The L.F.S. is a stratified multi-stage design employing probability sampling at all stages of the design. Dwellings from four rotation groups were specifically selected for the test. The Labour Force Survey was not conducted in dwellings selected independently for the test in order to avoid any possible respondent conditioning effects that might be introduced as a consequence of having participated in the L.F.S. In total, 39,889 dwellings were selected for use in the test. Of these dwellings, 4,432 were used for a special field edit response study. An additional 495 dwellings were added during the field collection operations, and questionnaires from 1,107 of the dwellings were not received or traced. The result was that a questionnaire with a legitimate response code was received from 34,845 dwellings. The following table shows an evaluation of the responding households for the National Census Test.

Responding households	26,651
Non responding Households	3,279
Out of Scope Dwellings*	4,956
Invalid final response codes	45
Total Dwellings	34,845

\* Defined as vacant, demolished, etc.

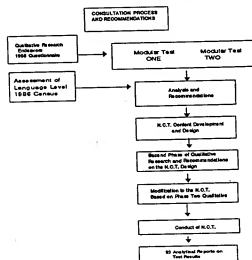
In the 26,651 responding households, information was collected for 71,596 persons. The response rate for the National Census Test based upon in scope dwellings was 60.3% prior to field follow-up, with the final response rate of 89.1% after field follow-up.

As was stated earlier, the N.C.T. was not intended as a dress rehearsal for the 1991 Census, however, the collection methodology did attempt to follow, to the degree possible, the collection procedures to be used for the 1991 Census. That is to say, the questionnaires were dropped off by enumerators during the week before "Census Test Day", which was November 4, 1989. Respondents were left a questionnaire package in the official language of their choice as well as a postage paid return envelope and a guide on how to complete the questionnaire. Households which had not mailed back their questionnaire by the second week of November were either followed-up by telephone or contacted by personal visit. As is the case in the census collection operations, the completed questionnaires were subjected to a series of field edits and quality control checks.

### 4. Choice of Approach for the Methodology

Given the objectives of the N.C.T., a number of decisions with respect to how the test would be conducted had to be made in the very early stages of the testing process. Statistics Canada was faced with the option of conducting one test on a large sample or with testing several versions of a questionnaire on smaller samples. Given a fixed budget and fairly tight time schedules the trade offs were assessed as follows. Using a large sample size on a single test provided the ability to conduct rather detailed analysis on estimates at provincial levels of geography and maintain fairly low sampling variability. In essence, the data from a single test questionnaire with a large sample size would permit the analysis of the results at a fairly detailed level of geography with scope for significant cross classifications of data elements. The other option of conducting the test with several versions of the questionnaire on smaller sample sizes would enable the testing of several order and wording formulations for the questions, but would leave the resulting data products with rather small samples and little, if any, scope for detailed analysis at the province level. Given that the primary objective of the test was intended to detect response error it was decided to conduct the test on the larger sample thereby permitting detailed cross classifications of tabulations at the level of province. Further, the analysis of the test data could be compared to 1986 Census estimates as well as data from administrative sources, or data from more recent sample surveys. With the large sample size options analysts could be fairly certain that any observed differences in data from the test and other sources were due to actual response differences and not sampling variability. It was for these reasons that the N.C.T. was conducted using a single version of a most likely 1991 census questionnaire on a sample of almost 40,000 dwellings.

### 5. The Model for the Development of the National Census Test



The diagram above illustrates the interplay between the elements which had significant influence in the development design and the content for the national Census Test. Many of the "input" factors to the test were either completed or in the process of finalization at the time that the N.C.T. project team began work on the test. The input elements provided their products and recommendations during a period that

Page 101

Page 102

Page 103

Page 104

Page 105

Page 106

extended over almost two years prior to the actual conduct of the test.

#### 6. Detection and Evaluation of Response Errors/Analytical Techniques

Given that the National Census Test was a probability sample, the production of weighted estimates was one of the key elements in the design of the data analysis portion of the testing program. The weighted estimates from the test were also accompanied by the production of exact variance measures so that analysts could make statements about the N.C.T. data tabulations with full confidence that the measures they were assessing were due to changes over time, changes in reporting patterns within the questionnaire, or response error.

The second method utilized to detect response error in the test questionnaire was the examination of edit failures. Within the N.C.T. sample a random selection of 4,432 dwellings were identified to be included in an edit failure study. This study was intended to assess reporting errors in questionnaires which had not been subjected to field edits or data processing edits in head office. Questionnaires for the response error study from these pre-identified dwellings were captured and analyzed exactly as completed by the responding household. The analysis from the response error study provided an excellent source for identification of questions with high non-response patterns or questions which exhibited a high proportion of incorrect reporting; for example, missed skip patterns, multiple responses, partial responses, etc. In most cases the results of the response error study indicated that the N.C.T. questions displayed lower respondent error rates than their counterparts in the 1986 Census. This response rate analysis provided a clear indication that many of the changes in the layout and design incorporated in the N.C.T. questionnaire had met the objective of providing respondents a more "user friendly" Questionnaire.

The third analytical technique used in the of evaluation and detection of response error in the N.C.T. was a "re-interview" of a ten percent sample of the households in the N.C.T. The objective of the reinterview was an attempt to measure response inconsistency. The reinterview was conducted approximately two to three weeks after the completion of the Census test. This reinterview aspect of the test proved to be the least conclusive with respect to the detection of response errors. While the reinterview identified a certain amount of response inconsistency between the N.C.T. questionnaire and the reinterview, it was not at all clear as to what accounted for the reporting inconsistency between the two. The reporting inconsistency between the N.C.T. results and the reinterview questionnaire had three possible sources. The first of these was proxy reporting. In the N.C.T., proxy reporting was permitted while in the reinterview no proxy reporting was allowed. It was felt that this factor accounted for some of the differences observed in the reporting patterns in the two questionnaires. The second difference between the N.C.T. and the reinterview questionnaire was the method of enumeration. In the case of the N.C.T., the questionnaire was a self-enumerated style of document. The reinterview, however, was administered by an interviewer either over the telephone or in person. These differences in the data collection methodology also accounted for a number of the reporting inconsistencies identified. Finally, the time lag between the conduct of the National Census Test and the reinterview was also deemed to be a source of reporting differences between the test and the reinterview. In short, it

was found that the interaction of the factors described above made the interpretation of the results of the reinterview questionnaire difficult, if not impossible, to assess. While the reinterview questionnaire contributed in a positive fashion for the assessment and evaluation of coverage questions, the reinterview data was not particularly helpful in the evaluation of reporting discrepancies for most of the other variables on the N.C.T. questionnaire.

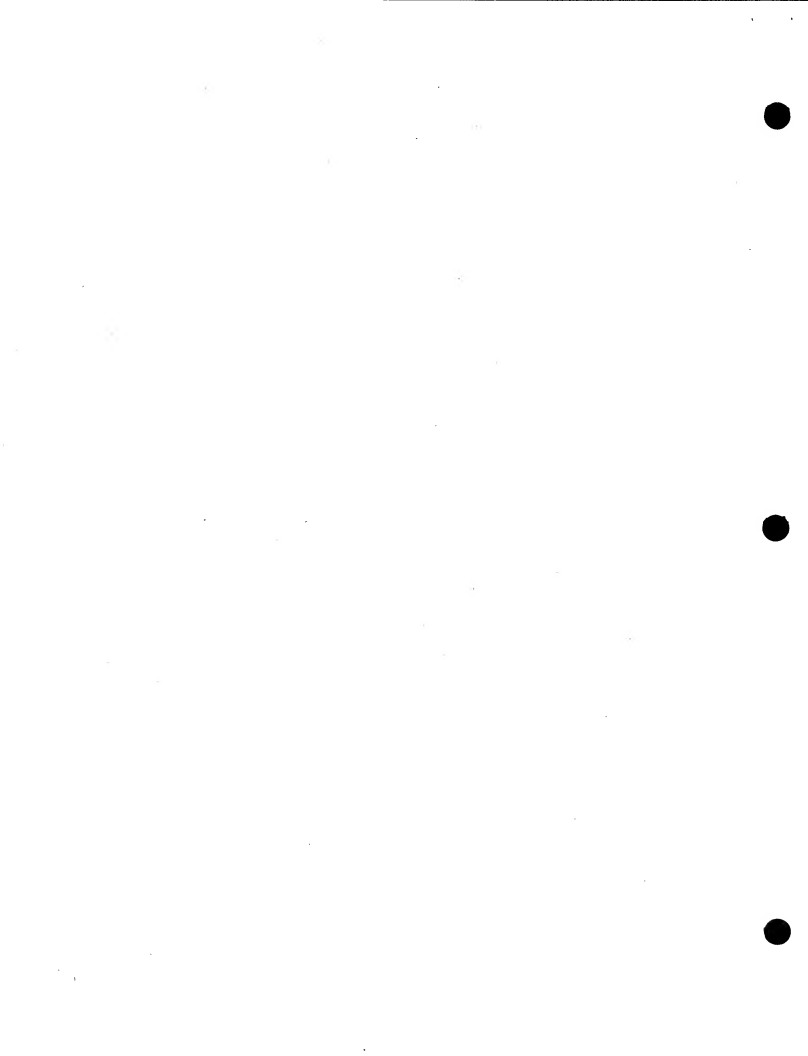
#### 7. Attempted Questionnaire Design Improvements

The primary inputs to the questionnaire design changes for the National Census Test were predicated on the results of the qualitative research conducted on the 1986 Census questionnaire. With a test environment and few limiting constraints on the design of the questionnaire the N.C.T. project team started the design of the test instrument with virtually "a clean slate".

The major change in the design of the National Census Test questionnaire was the move from the 1986 Census questionnaire to a matrix format questionnaire. This represented one of the most significant overall changes in the design of the test questionnaire. The form produced for the test utilized a design layout which permitted all household members (up to six persons) to enter their names across the top of the form and complete blocks of questions for all members of the household at the same time, or respond on the basis of a single person at a time. The major concern with the introduction of a matrix format was that respondents would not transcribe the names of household members in the order they appeared in the roster question. The analysis of the test results confirmed that the matrix approach worked well and that the respondents ability to cope with a matrix format was, if anything, as good as or better than the 1986 Census design layout of the questionnaire. The matrix layout used in the National Census Test was recommended for use in the 1991 Census questionnaire.

The National Census Test also introduced the use of a roster question. The intent of the roster question (Step 2) was to let respondents list all of the members of the household prior to actually starting the responses to the questionnaire. The roster question was introduced as a coverage improvement measure for the test. This question was further reinforced by the addition of a supplementary inquiry as to whether anyone had been left out of the roster listing because the respondent was unsure if the person should have been included. The analysis of the roster question indicated that it worked very well, despite a number of instructions as to whom to include in the roster as well as how the persons should be listed on the roster. As a result of the analysis of the test it was recommended that the roster and associated questions be retained for the 1991 Census of Canada.

The results of the qualitative research, as well as the debriefing of interviewers who were involved in the conduct of the test, strongly supported the widely held view that respondents do not read instructions. In the design of the N.C.T. questionnaire an effort was made to either eliminate instructions by improvements to the questions, or significantly reduce the verbiage where instructions were provided. In general, this attempt to introduce "white space" in the design of the document was successful. Many of the instructions which were expunged from the questionnaire eventually found their place in the respondents guide.



While respondents may gloss over instructions, they seem to dedicate their time to at least reading the questions and would seem to peruse response categories very carefully. The following two examples, from the analysis of questions on language abilities with either of Canada's two official languages and questions on ethnic origin and identity, best illustrate the point.

In the 1986 Census of Canada a question on language abilities appeared in the long version of the Census form. That is the form 2B which is distributed on a sample basis to every fifth dwelling. The language question asked on the 1986 Census form was:

**Can you speak English or French well enough to conduct a conversation?**

As a result of the consultation process and the modular tests, it was decided that perhaps the criteria of what constituted the ability to speak could be better defined within the question per se. The result was a modification in the language ability question which was tested in the N.C.T. The question used in the test was as follows:

**What language or languages does this person speak well enough to conduct a fairly lengthy conversation on different topics?**

The N.C.T. test questionnaire provided a more stringent approach than the 1986 Census question. Although for all intents and purposes the questions were intended to measure the same basic phenomena. The results, however, as shown graphically in Table 1 clearly indicate statistically significant differences between the two measures. The results of the analysis of this question lead to the conclusion that respondents were in fact reading the details of the questions and in general seem to interpret the question in the appropriate manner. The comparison of the results from the 1986 Census and the N.C.T. show a decrease in respondents' abilities to speak both of Canada's two official languages. Thus, the strengthened criteria and threshold, as established in the test questionnaire, was correctly interpreted by respondents. The results from this question were further corroborated with comparisons of the results of a detailed household survey conducted by Statistics Canada as part of the General Social Survey program. The questions on this Social Survey were very detailed with respect to language and permitted an in-depth comparison of the results of the N.C.T. language question results. The findings from the N.C.T. were found to be comparable with the detailed language questions in the 1986 General Social Survey (G.S.S.) which used a set of 8 to 10 questions to define capabilities in both official languages. Table 2 presents the results of the 1986 G.S.S. data on ability to conduct a conversation in English, French or another language and the results of the National Census Test question which was intended to measure the same characteristic.

The second example of the analysis of respondent's ability to read and comprehend the questions as developed for the N.C.T. questionnaire, is demonstrated by the questions on ethnic origin and identity. These are two very complex areas of investigation in survey research because of the subjective or "soft" nature of the topic, as well as multiplicity of possible measures suggested by researchers in the field as "the correct one". Question 15 and 16 of the National Census Test were intended to collect data on ethnic origin and identity. Several possible options for questions to provide a measure on this phenomena were discussed with numerous experts in the field.

In addition, the results of the qualitative analysis and Modular Test 2 brought to bear on the development of these two questions. After much debate amongst "experts" in the area the decision was made to use a question which would get a measure of "ancestry" for origin information, without actually using the term "ancestry". The principal reason for this decision being that focus group testing determined that the word "ancestry" was not well-understood by respondents. The terms used to define the question on ancestry were as follows:

**What are the ethnic or Cultural origins of this persons  
Parents and grandparents?**

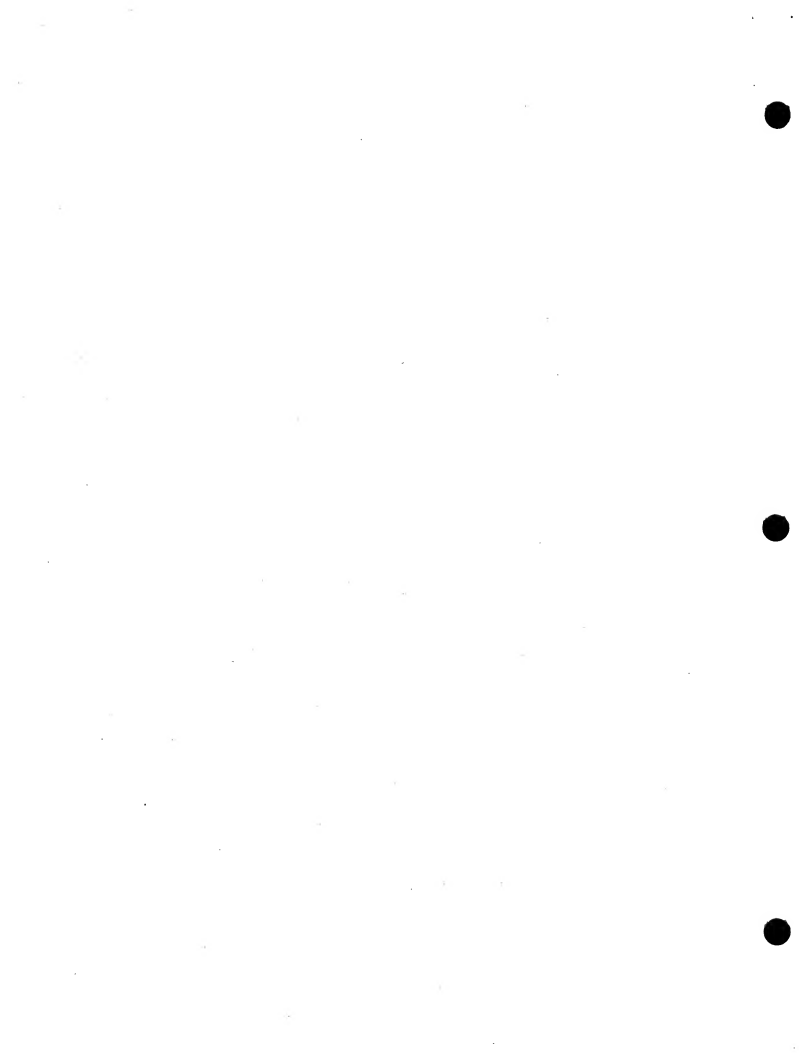
The question used for identity was:

**What is this persons ethnic or cultural identity?**

The results presented in Table 3 clearly indicate that respondents, despite being confronted with a set of response categories which were identical for each of these questions showed, different reporting patterns on the questions on ancestry and identity. The detailed analysis of the data firstly pointed out that respondents did appear to differentiate between the two concepts. Further, it is clear from the results of the analysis of the two questions that respondents saw their identity as being different from their ancestry. This was evidenced in the different reporting patterns in the parents and grandparents (ancestors) questions. Analysis of the results showed that respondents review all closed ended response category options before making their selection. This is evidenced by the fact that the response category for "Canadian" in the questions was the last in a list of 16 possibilities for a response in both the questions on ancestry and identity. The graphic illustration in Table 4 clearly demonstrates that respondents "found" the Canadian category amongst all of the other possible response options. Sixteen percent of respondents defined their ancestry as "Canadian" and 37.2% defined their identity as Canadian. In the 1986 Census, where respondents had to "write in" or "specify" Canadian as a response, less than half a percent of the population did so. The analytical techniques described in the two selected examples above were utilized for all variables on the questionnaire. The end result was the preparation of 23 analytical reports on the N.C.T., which were utilized by Statistics Canada's senior management in the decision-making process of what would constitute the content of the 1991 Census of Canada.

One of the major criticisms of the 1986 Census document was that it had no definite start point. For the N.C.T., a "step" approach was developed for the questionnaire design. The respondents were told to start at step one and to continue their way through the questionnaire by following the steps. In testing the step approach in focus groups, respondents found that its use was a significant improvement over the 1986 Census form and that the steps were easy to follow. In addition to giving the N.C.T. questionnaire a definite start point, respondents were also given the A,B,C's of how to complete the document. The directions on how to complete the form were placed at the front of the questionnaire and instructed the respondent on how to fill in the form, where to get help in filling out the form and what to do with the questionnaire once it was completed. While these sound like very obvious instructions they were not easily found on the 1986 Census form.

Significant changes were made to the print size in the N.C.T. questionnaire. Print fonts two sizes larger than the 1986





Census were used. In addition, the mark boxes used in 1986 were changed to circles for the N.C.T. so that respondents were asked to either check a response circle, enter a number, or print in a response box. On the 1986 Census form, response categories were made up of check-boxes or write-in boxes which caused some confusion on the part of respondents.

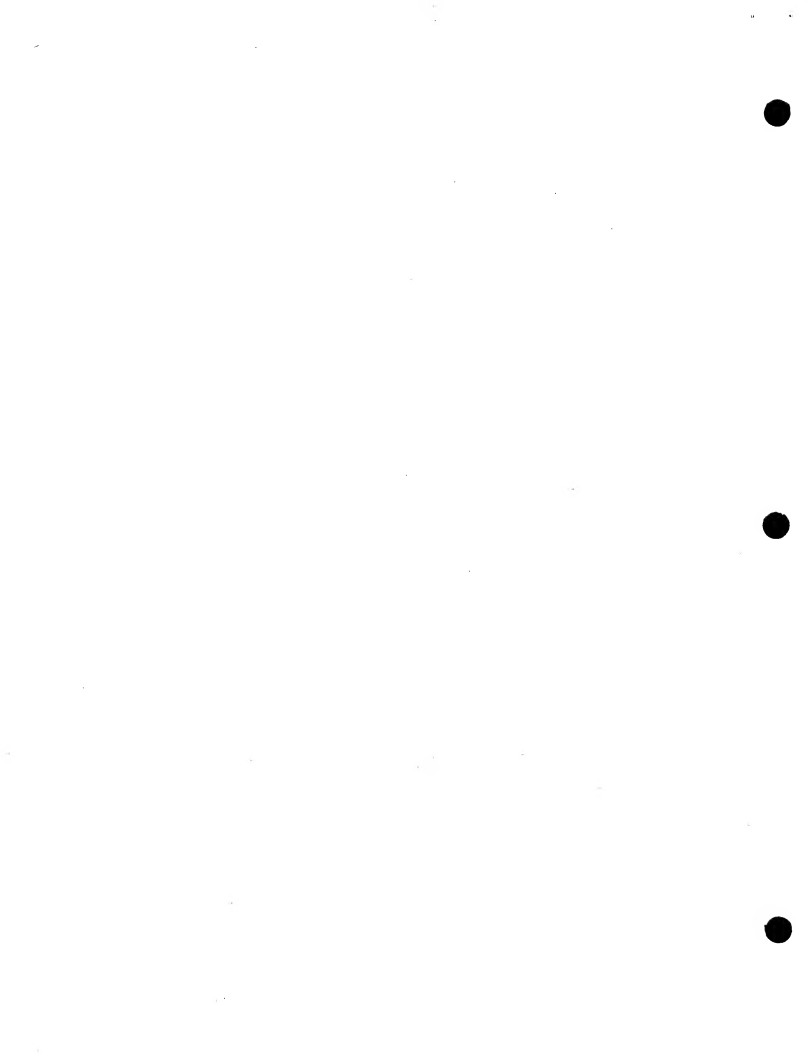
The 1986 Census long form questionnaire contained a fold-out flap on the first page of the document. This fold-out flap contained a message to the respondents as well as some very detailed instructions on how to complete the first question on the form. The flap caused a number of problems for the respondent in terms of handling the questionnaire, as well as in the ensuing data processing operations. With the movement in the N.C.T. to a reduction in both the number and the verbiage in instructions the flap was eliminated from the test questionnaire.

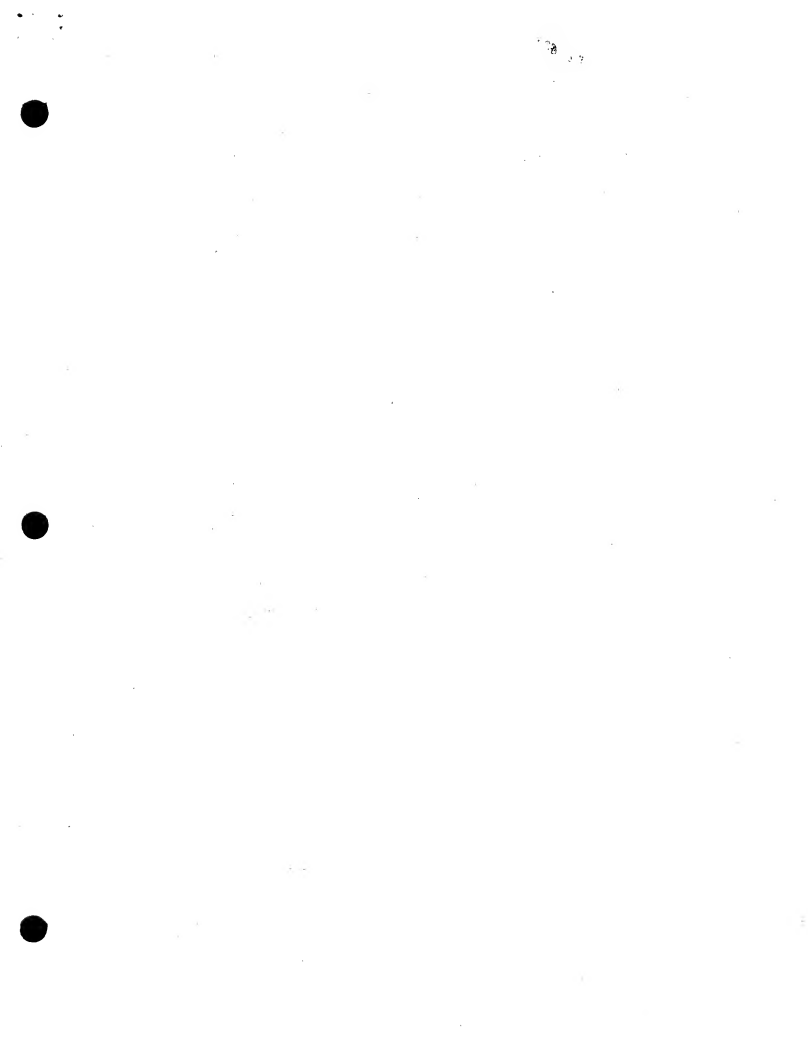
One of the criticisms leveled against the 1986 questionnaire was the fact that there was no readily apparent "end point" to the document. While this seems a petty point, from a respondent relation point of view, it is an important factor in fostering goodwill. Modifications to the N.C.T. questionnaire included a definite end point, a thank you and a note that the

respondent should mail the completed questionnaire as soon as possible.

#### 8. Conclusions

The results of the National Census Test analysis were made available in a series of 23 individual analytical reports. These reports provided analysis of the major question modules (e.g. demography, coverage, language, mobility, labour market, income, etc.). In general, the conclusions from the analytical endeavors provided strong evidence that the attempted improvements in the restructure and redesign of the questionnaire were productive. Many of the changes suggested in the test program were carried forward in the design of the 1991 Census of Canada. As in the case of a large and complex undertaking like the Census, some of the improvements which were operationally feasible in a test setting were not transferable to the large scale operational activities associated with the census collection and processing projects.





CS 018

STATISTICS CANADA LIBRARY  
BIBLIOTHEQUE STATISTIQUE CANADA



1010197450